

Release Notes

DOI: [10.15154/z563-zd24](#) (Release 5.1)

PUBLISHED
August 5, 2026

General Information

An overview of the ABCD Study® can be found at [abcdstudy.org](#) and detailed descriptions of the assessment protocols are available at [ABCD Protocols](#). This page provides important general information about the ABCD 5.1 data release.

Sample

Release 5.1 contains tabulated data and minimally processed imaging data for visits started between September 1, 2016 and January 15, 2022. It includes data from 11,868 participants, the entire ABCD cohort (N = 11,880) except for twelve participants who withdrew consent to share their data, and nine events. Events up to the 3-year follow-up event are completed, with varying numbers of missed visits per event; the 42-month and 4-year follow-up events were still ongoing when the data for this release was frozen, so data are included for a subset of participants for these timepoints.

The following table lists the number of included participants per event:

Event	n
baseline	11868
6-month follow-up	11389
1-year follow-up	11220
18-month follow-up	11083
2-year follow-up	10973
30-month follow-up	10228
3-year follow-up	10336
42-month follow-up	8449
4-year follow-up	4754

5.1 Release Data Changes

The 5.1 Patch Release addressed some issues from the 5.0 Data Release. There are some discrepancies between the 5.0 and 5.1 Data, due to changes that were made, but 5.1 Data is the most up to date. Here are the updates for 5.1:

Imaging

- The "mri_y_qc_incl" file did not include 4 year follow up data. The file has been updated with the 4 year follow up data.
- In the MID behavioral dataset, the variables "tfmri_mid_all_beh_hrw_mrt" and "tfmri_mid_all_beh_nt_mrt" had identical values for every row of the data set. This issue has been fixed.
- 4 year events were missing from the MRI raw QC tables. This issue has been fixed.
- There is an issue with missing mproc data. The pGUIDS that have been affected are listed in the knownissue_5.1_mproc zip file. This issue will be addressed in Release 6.0.

Non-imaging

- The KSADS summary scores had missing data. The data have been updated. There are still issues with the diagnoses for eating disorders, so please use the symptoms for analyses. Please check [here](#) for more info. The previous 5.0 KSADS raw data also had some issues with swapped pGUIDS (e.g. parents/guardians who mistakenly answered questions for one child when they meant to respond about the other)
- These summary scores had missing data and have been updated in 5.1 release:
 - CBCL
 - Brief Problem Monitor
 - 7-Up Mania inventory (mh_y_7up)
- For the Resilience Measure, the data frame on R has mis-specified values, due to open-ended responses. We have applied the following criteria (e.g. "none", "no" -> 0, max value of 100, average of 2 numbers and rounded down). The data have been updated.
- Some of the "meim-r" variables in the ce_p_meim table are incorrect. The data have been updated.
- Some of the site IDs were incorrect and have been corrected.

The following tables have changes:

domain	table	field(s)	issue
ABCD (General)	abcd_y_lt	site_id_l	Incorrect values
Culture & Environment	ce_p_meim	meim_[1-6]f_p	Incorrect values
Imaging	mri_y_adm_nts.csv	multiple	Missing/incorrect values

domain	table	field(s)	issue
Imaging	mri_y_adm_qtn.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_clfind.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_incl.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_raw_dmr.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_raw_rsfmt.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_raw_smr_t1.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_raw_smr_t2.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_raw_tfmr_all.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_raw_tfmr_mid.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_raw_tfmr_nback.csv	multiple	Missing/incorrect values
Imaging	mri_y_qc_raw_tfmr_sst.csv	multiple	Missing/incorrect values
Imaging	mri_y_tfmr_mid_beh.csv	multiple	Missing/incorrect values
Imaging	mri_y_tfmr_nback_beh.csv	multiple	Missing/incorrect values
Imaging	mri_y_tfmr_nback_rec_beh.csv	multiple	Missing/incorrect values
Imaging	mri_y_tfmr_sst_beh.csv	multiple	Missing/incorrect values
Mental Health	mh_p_cbcl	summary scores	Missing/incorrect values
Mental Health	mh_y_bpm	summary scores	Missing values
Mental Health	mh_t_bpm	summary scores	Missing values
Mental Health	mh_y_7up	summary scores	Missing values
Mental Health	mh_y_or	resiliency[5-7]a_y , resiliency7b_y	Invalid (string instead of numeric or out of range) responses
Mental Health	mh_p_ksads_ss.csv	symptom & diagnosis scores	Missing/incorrect values
Mental Health	mh_y_ksads_ss.csv	symptom & diagnosis scores	Missing/incorrect values

domain	table	field(s)	issue
Mental Health	mh_p_ksads_adhd.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_ago.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_asd.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_bg.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_bp.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_cd.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_dep.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_dmdd.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_ed.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_gad.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_hi.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads OCD.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_odd.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_pd.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_phb.csv	multiple	Incorrect/switched participant IDs

domain	table	field(s)	issue
Mental Health	mh_p_ksads_psy.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_ptsd.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_sad.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_sep.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_si.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_p_ksads_slp.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_y_ksads_bg.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_y_ksads_bip.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_y_ksads_cd.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_y_ksads_dep.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_y_ksads_dmdd.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_y_ksads_ed.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_y_ksads_gad.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_y_ksads_sad.csv	multiple	Incorrect/switched participant IDs
Mental Health	mh_y_ksads_si.csv	multiple	Incorrect/switched participant IDs

domain	table	field(s)	issue
Mental Health	mh_y_ksads_slp.csv	multiple	Incorrect/switched participant IDs
Substance Use	su_y_ksads_sud.csv	multiple	Incorrect/switched participant IDs
Substance Use	su_p_ksads_sud.csv	multiple	Incorrect/switched participant IDs
Substance Use	su_y_hair_tox	su_y_hair_selection	Missing values

Significant Changes for 5.0

Data Dictionary Explorer Application

The 5.0 release introduces a new data dictionary explorer application (<https://data-dict.abcdstudy.org/>) that replaces the static data dictionary functionality on NDA's website. The application contains all relevant information about the variables in the ABCD Study core protocol as well as substudies in a single resource. It allows exploration of the ontology of ABCD instruments, i.e., their hierarchical organization into domains and subdomains, in an interactive manner and provides extensive filter and search capabilities.

The data dictionary lists variables' names, labels (descriptions), the table they belong to, and additional information like branching/skip logic. For categorical variables, it lists the numeric values and the labels they refer to (e.g., 0 = "No"; 1 = "Yes"; 999 = "Don't know"). To learn more about this application and how to use it to explore the ABCD data resource, see [Data Dictionary Explorer](#).

Online Release Notes

Beginning with the 5.0 release, release notes are published on the ABCD Study wiki. These notes contain information about the data collection instruments for all ABCD domains and substudies, including a description of the measure, relevant background information and references, known issues, and potential modifications of the instrument over the course of the study. Users should carefully read and consider this information when analyzing data from the ABCD Study. Users interested in ABCD's imaging data should also refer to the [Imaging Overview](#) release note.

The new online format allows ABCD to update the release notes more frequently. ABCD will work to continuously improve the resource, include additional information, and will likely add more features like an FAQ section.

Please regularly check the website for updates!

Notes:

- The website provides full-text search functionality; to search for terms of interest over all pages, click on the magnifying glass on the right side of the menu bar.
- The *Changes and Known Issues* release note contains information about individual cases and requires that users have a valid data use certificate. It is available on the ABCD [study page](#) on the NDA website. We will notify users of any changes to this release note on this website and recommend that users download the most up-to-date version from NDA.

Naming and structure of data tables

ABCD has started a process to curate the data resource in a more consistent and standardized manner. As a first step, ABCD developed a standard for naming data tables that was implemented in the 5.0 release. This standard takes into account two hierarchy levels that every table can be assigned to:

Domain

- *ABCD Core*
 - ABCD (General): **abcd**
 - Culture & Environment: **ce**
 - Gender Identity & Sexual Health: **gish**
 - Genetics: **gen**
 - Imaging: **mri**
 - Linked External Data: **led**
 - Mental Health: **mh**
 - Neurocognition: **nc**
 - Novel Technologies: **nt**
 - Physical Health: **ph**
 - Substance Use: **su**
- *ABCD Substudies*
 - COVID-19: **cvd**
 - Endocannabinoid: **ecb**
 - Hurricane Irma: **irma**
 - Social Development: **sd**

Source

- Youth: **y**
- Parent: **p**
- Teacher: **t**
- Linked Dataset: **l**

These levels are used as a prefix and are combined with an acronym for the instrument to give the standardized table name, for example:

Domain	Source	Instrument	Table Name
Mental Health	Parent	Child Behavior Checklist	mh_p_cbc1
Culture & Environment	Youth	Parental Monitoring	ce_y_pm
Culture & Environment	Parent	Parental Monitoring	ce_p_pm
Linked External Data	Linked Dataset	Building Density	led_l_densbld
...	...	...	...

Note: The Imaging domain takes into account additional hierarchy levels for their table names which is explained in more detail in the [Imaging release notes](#).

In addition to changing table names, ABCD also reconsidered the way in which variables are grouped into different tables. On the one hand, this involved breaking up tables that previously contained a large number of variables into several tables with a smaller, more coherent set of variables (e.g., in the [Imaging](#) domain). On the other hand, it involved combining variables with consistent content from previously separate tables into one table, especially,

- individual items ("raw data") of an instrument are now combined with the summary scores for that instrument
- variables from the baseline version of an instrument are now combined with the longitudinal version of that instrument (see [Baseline and Longitudinal Instruments](#) below).

The accompanying [data dictionary explorer](#) application allows users to explore the structure of the ABCD data resource in an interactive manner (see [above](#)). It illustrates how variables are grouped into the new tables and how tables are hierarchically organized within the ABCD ontology. For backward-compatibility, the application also allows users to look up the table names as well as DEAP variable names used in the 4.0 release. As such, users can search for a variable name or a 4.0 table name in the interactive data dictionary table and see which table this variable now belongs to (for more information on how to use the application, see [Data Dictionary Explorer](#)).

ABCD hopes that the standardization and reorganization of tables will improve the overall user experience and make it easier to find related variables and constructs in the dataset. As a next step in the efforts to make the data resource more consistent and standardized and as an extension of the new table naming scheme, ABCD plans to develop a standardized naming scheme for all variables in the resource.

Tabulated Release Data

File type and structure

In release 5.0, the tabulated data is not provided through NDA's Oracle database. Instead of downloading it using the NDA Download Manager, data users will be able to download the tabulated

data from the ABCD [study page](#) on the NDA website.

All tabulated data files are provided as one zip archive file with approximately 5 GB in size (15 GB after extraction). After extracting the archive, the root directory will contain several subdirectories grouping the data files by domain. The tables are provided as plain text files (`.csv`) that can be imported into any statistical software for analysis.

Notes

- In contrast to previous releases, files are provided as *comma-separated* instead of *tab-separated* text files.
- In contrast to previous releases, files do not contain an additional header line, i.e., they have a more standard format with *one header row* listing the variable names and *data starting in row two*.

Data format

There are some notable changes to previous releases in how the data is exported and structured. These include

- NDA's standard fields are not included in every table anymore
 - Files from previous releases contained an NDA-specific selection of standard fields like sex at birth, assessment age, assessment date, etc. *in every file*.
 - In contrast, the new files only contain the two key columns (`src_subject_id` and `eventname`) which can be used to join any other variables of interest from other tables
- Each table only includes a row for a given participant/event if there was data for at least one of the variables included in the table, resulting in a sparser representation of the data.

Export of checkbox items

Probably the most relevant data-related deviation from previous releases is the way in which checkbox fields (multi-select or "select all that apply" questions) are exported from REDCap.

Generally, checkbox items are exported in the following manner

- Each choice option for a given checkbox field becomes their own `0/1` (yes/no) variable in the exported data
- For example, a checkbox field `chkb` with 5 multiple choice options becomes 5 fields
 - `chkb__1`
 - `chkb__2`
 - ...
 - `chkb__5`

In previous releases, the standard API method for exporting records from REDCap was used. This resulted in the following output format:

- If a participant *selected one or several options*,
 - the variables for these options would have the *value 1*
 - the variables for the other options would have the *value 0*

- If a participant did *not select any of the options* (independent of whether they missed the specific item or did not even receive the respective survey at all),
 - all variables associated with the checkbox item would have the *value 0*

Case	chkb__1	chkb__2	chkb__3	chkb__4	chkb__5
One selection (2)	0	1	0	0	0
Two selections (3, 5)	0	0	1	0	1
No selections	0	0	0	0	0

This is an unexpected and potentially misleading behavior because it suggests that the data is based on actual participant behavior/choices. Especially when users only consider one of the response options for their analysis, the old format might result in misinterpretation and wrongful inclusion of cases. In the 5.0 release, we decided to export the data using an alternative method which results in the following output format:

- not changed: If a participant *selected one or several options*,
 - the variables for these options would have the *value 1*
 - the variables for the other options would have the *value 0*
- changed: If a participant *did not select any of the options*,
 - all variables associated with the checkbox item are set to a *missing value*

Case	chkb__1	chkb__2	chkb__3	chkb__4	chkb__5
One selection (2)	0	1	0	0	0
Two selections (3, 5)	0	0	1	0	1
No selections					

Additional Information

COVID-19 and In-person, Remote, and Hybrid Testing

In response to COVID-19 restrictions beginning in March 2020, ABCD pivoted to remote testing when in-person testing was not possible or feasible and subsequently a hybrid in-person/remote testing procedure as sites allowed. This affects all annual assessments conducted in March 2020 or later (in the 5.0 release, those are 2-, 3, and 4-year follow-up visits).

Remote and hybrid testing required participants to complete some tasks and surveys on their own devices (i.e., phone, tablet, desktop, or laptop computer). Note that remote performance was monitored by research associates, when possible, using Zoom's screen sharing feature. The variety of devices, relative to the ABCD standard using Apple iPad devices exclusively, may affect task performances and users should consider this when analyzing data spanning the pre-COVID-19 and post-COVID-19 periods. In addition, some tasks were incompatible with remote testing and were not administered during this time. Please refer to the release notes of the different assessment domains for guidance on how the remote procedures might affect participants' behavior or responses.

Note: To determine the visit type for a given participant/event, e.g., to account for it in statistical analyses, the *Longitudinal Tracking* instrument ([abcd_y_lt](#)) includes the variable `visit_type` that codes the assessment setting in the following manner:

- 1 = In person
- 2 = Remote
- 3 = Hybrid

Spanish Versions of Parent Instruments

Parent questionnaires used by ABCD are dual language (English and Spanish). Parents can choose to display the instrument in Spanish by selecting a checkbox displayed at the top of every instrument:

¿Español? — ☒ Sí

This allows the ABCD project to use a single instrument to capture responses for both languages (see [redcap-hook-framework](#) on the ABCD GitHub page).

We are sharing the Spanish language descriptions of items, but due to technical limitations of the export process (no support for UTF-8) the Spanish language content appeared broken in the NDA data dictionaries. As the 5.0 data dictionary is based on these data dictionaries, this is still the case. We are planning to fix the Spanish content in the [data dictionary explorer application](#). For the meantime, we hope that the visible content is sufficient to inform the reader about the Spanish wording of questions. Please [contact](#) the ABCD Coordinating Center (CC) if you require access to the original Spanish language version of REDCap instruments.

Baseline and Longitudinal Instruments

Several instruments ask similar questions at the baseline and follow-up visits. For example, in the [Substance Use Interview](#), ABCD captured data in the following manner

- Baseline: "Have you ever tried drug x at any time in your life?"
- 1-year follow-up: "Have you used any of the following drugs *since we last saw you on {date}*?—Drug x?"

Due to the different time frame and wording of the baseline and follow-up questions, distinct variables were used to capture these data. To indicate the association, the suffix `_l` was added to the baseline variable name to create the longitudinal variable name used in follow-up events (e.g., a baseline variable named `drug_used` would be named `drug_used_l` as a longitudinal variable). While baseline and longitudinal variables were originally shared in separate tables, they were combined in the new table structure (see [above](#)) to make the association more obvious to users and facilitate analysis.

Survey items: Directions and Descriptions

Many instruments contain descriptive text and instructions used during data collection as prompts to the participant/parent. In some cases, these descriptive texts that sometimes precede multiple items define the meaning of the responses. They are typically not included in the ABCD data dictionary because only variables that refer to actual data values, i.e., columns in the data tables, should be included. ABCD is planning to improve the data dictionary explorer application in future releases to display this important meta information.

Known Issues

Following is a list of known issues from ABCD 5.1 Data Release:

Non-Imaging

- There are remaining issues with the KSADS Eating Disorder diagnoses. For more info, please refer to [here](#).

Imaging

- There is an issue with missing mproc data. The pGUIDS that have been affected are listed in the `knownissue_5.1_mproc` zip file. This issue will be addressed in Release 6.0